

# The number of triple systems without even cycles

Dhruv Mubayi\*      Lujia Wang†

February 10, 2017

## Abstract

For  $k \geq 4$ , a loose  $k$ -cycle  $C_k$  is a hypergraph with distinct edges  $e_1, e_2, \dots, e_k$  such that consecutive edges (modulo  $k$ ) intersect in exactly one vertex and all other pairs of edges are disjoint. Our main result is that for every even integer  $k \geq 4$ , there exists  $c > 0$  such that the number of triple systems with vertex set  $[n]$  containing no  $C_k$  is at most  $2^{cn^2}$ .

An easy construction shows that the exponent is sharp in order of magnitude. This may be viewed as a hypergraph extension of the work of Morris and Saxton, who proved the analogous result for graphs which was a longstanding problem. For  $r$ -uniform hypergraphs with  $r > 3$ , we improve the trivial upper bound but fall short of obtaining the order of magnitude in the exponent, which we conjecture is  $n^{r-1}$ .

Our proof method is different than that used for most recent results of a similar flavor about enumerating discrete structures, since it does not use hypergraph containers. One novel ingredient is the use of some (new) quantitative estimates for an asymmetric version of the bipartite canonical Ramsey theorem.

## 1 Introduction

An important theme in combinatorics is the enumeration of discrete structures that have certain properties. Within extremal combinatorics, one of the first influential results of this type is the Erdős-Kleitman-Rothschild theorem [25], which implies that the number of triangle-free graphs with vertex set  $[n]$  is  $2^{n^2/4+o(n^2)}$ . This has resulted in a great deal of work on problems about counting the number of graphs with other forbidden subgraphs [6, 7, 8, 14, 15, 26, 31, 40, 48] as well as similar question for other discrete structures [10, 11, 17, 18, 35, 46, 47, 49, 51]. In extremal graph theory, these results show that such problems are closely related to the corresponding extremal problems. More precisely, the *Turán* problem asks for the maximum number of edges in a (hyper)graph that does not contain a specific subgraph. For a given  $r$ -uniform hypergraph (henceforth  $r$ -graph)  $F$ , let  $ex_r(n, F)$  be the maximum number of edges among all  $r$ -graphs  $G$  on  $n$  vertices that contain no copy of

---

\*Department of Mathematics, Statistics, and Computer Science, University of Illinois, Chicago, IL 60607, Email:mubayi@uic.edu. Research partially supported by NSF grant DMS-1300138.

†Department of Mathematics, Statistics, and Computer Science, University of Illinois, Chicago, IL 60607, Email:lwang203@uic.edu. Research partially supported by NSF grant DMS-1300138.

$F$  as a (not necessarily induced) subgraph. Henceforth we will call  $G$  an  $F$ -free  $r$ -graph. Write  $\text{Forb}_r(n, F)$  for the set of  $F$ -free  $r$ -graphs with vertex set  $[n]$ . Since each subgraph of an  $F$ -free  $r$ -graph is also  $F$ -free, we trivially obtain  $|\text{Forb}_r(n, F)| \geq 2^{\text{ex}_r(n, F)}$  by taking subgraphs of an  $F$ -free  $r$ -graph on  $[n]$  with the maximum number of edges. On the other hand for fixed  $r$  and  $F$ ,

$$|\text{Forb}_r(n, F)| \leq \sum_{i \leq \text{ex}_r(n, F)} \binom{\binom{n}{r}}{i} = 2^{O(\text{ex}_r(n, F) \log n)},$$

so the issue at hand is the factor  $\log n$  in the exponent. The work of Erdős-Kleitman-Rothschild [25] and Erdős-Frankl-Rödl [26] for graphs, and Nagle-Rödl-Schacht [45] for hypergraphs (see also [44] for the case  $r = 3$ ) improves the upper bound above to obtain

$$|\text{Forb}_r(n, F)| = 2^{\text{ex}_r(n, F) + o(n^r)}.$$

Although much work has been done to improve the exponent above (see [1, 6, 7, 8, 31, 34, 48] for graphs and [10, 11, 21, 47, 13, 50] for hypergraphs), this is a somewhat satisfactory state of affairs when  $\text{ex}_r(n, F) = \Omega(n^r)$  or  $F$  is not  $r$ -partite.

In the case of  $r$ -partite  $r$ -graphs, the corresponding questions appear to be more challenging since the tools used to address the case  $\text{ex}_r(n, F) = \Omega(n^r)$  like the regularity lemma are not applicable. The major open problem here when  $r = 2$  is to prove that

$$|\text{Forb}_r(n, F)| = 2^{O(\text{ex}_r(n, F))}.$$

The two cases that have received the most attention are for  $r = 2$  (graphs) and  $F = C_{2l}$  or  $F = K_{s,t}$ . Classical results of Bondy-Simonovits [16] and Kovári-Sós-Turán [36] yield  $\text{ex}_2(n, C_{2l}) = O(n^{1+1/l})$  and  $\text{ex}_2(n, K_{s,t}) = O(n^{2-1/s})$  for  $2 \leq s \leq t$ , respectively. Although it is widely believed that these upper bounds give the correct order of magnitude, this is not known in all cases. Hence the enumerative results sought in these two cases were

$$|\text{Forb}_2(n, C_{2l})| = 2^{O(n^{1+1/l})} \quad \text{and} \quad |\text{Forb}_2(n, K_{s,t})| = 2^{O(n^{2-1/s})}.$$

In 1982, Kleitman and Winston [32] proved that  $|\text{Forb}_2(n, C_4)| = 2^{O(n^{3/2})}$  which initiated a 30-year effort on searching for generalizations of the result to complete bipartite graphs and even cycles. Kleitman and Wilson [33] proved similar results for  $C_6$  and  $C_8$  in 1996 by reducing to the  $C_4$  case. Finally, Morris and Saxton [42] recently proved that  $|\text{Forb}_2(n, C_{2l})| = 2^{O(n^{1+1/l})}$  and Balogh and Samotij [14, 15] proved that  $|\text{Forb}_2(n, K_{s,t})| = 2^{O(n^{2-1/s})}$  for  $2 \leq s \leq t$ . Both these results used the hypergraph container method (developed independently by Saxton and Thomason [50], and by Balogh-Morris-Samotij [13]) which has proved to be a very powerful technique in extremal combinatorics. For example, [13] and [50] reproved  $|\text{Forb}_r(n, F)| = 2^{\text{ex}_r(n, F) + o(n^r)}$  using containers.

There are very few results in this area when  $r > 2$  and  $\text{ex}_r(n, F) = o(n^r)$ . The only cases solved so far are when  $F$  consists of just two edges that intersect in at least  $t$  vertices [9], or when  $F$  consists of three edges such that the union of the first two is equal to the third [12] (see also [4, 5, 22, 23] for some related results). These are natural points to begin these investigations since the corresponding extremal problems have been studied deeply.

Recently, Kostochka, the first author and Verstraëte [37, 38, 39], and independently, Füredi and Jiang [29] (see also [30]) determined the Turán number for several other families of  $r$ -graphs including paths, cycles, trees, and expansions of graphs. These hypergraph extremal problems have proved to be quite difficult, and include some longstanding conjectures. Guided and motivated by these recent developments on the extremal number of hypergraphs, we consider the corresponding enumeration problems focusing on the case of cycles.

**Definition 1** *For each integer  $k \geq 3$ , a  $k$ -cycle  $C_k$  is a hypergraph with distinct edges  $e_1, \dots, e_k$  and distinct vertices  $v_1, \dots, v_k$  such that  $e_i \cap e_{i+1} = \{v_i\}$  for all  $1 \leq i \leq k-1$ ,  $e_1 \cap e_k = \{v_k\}$  and  $e_i \cap e_j = \emptyset$  for all other pairs  $i, j$ .*

Sometimes we refer to  $C_k$  as a *loose* or *linear* cycle. To simplify notation, we will omit the parameter  $r$  when the cycle  $C_k$  is a subgraph of an  $r$ -graph.

Since  $\text{ex}_r(n, C_k) = O(n^{r-1})$ , we obtain the upper bound

$$|\text{Forb}_r(n, C_k)| = 2^{O(n^{r-1} \log n)}$$

when  $r$  and  $k$  are fixed and  $n \rightarrow \infty$ . Our main result is the following theorem, which improves this upper bound and generalizes the Morris-Saxton theorem [42] to 3-graphs.

**Theorem 2 (Main Result)** *For integers  $r, k \geq 3$ , there exists  $c = c(r, k)$ , such that*

$$|\text{Forb}_r(n, C_k)| < \begin{cases} 2^{cn^2} & \text{if } r = 3 \text{ and } k \text{ is even,} \\ 2^{cn^{r-1}(\log n)^{(r-3)/(r-2)}} & \text{if } r > 3. \end{cases}$$

Since trivially  $\text{ex}_r(n, C_k) = \Omega(n^{r-1})$  for all  $r \geq 3$ , we obtain  $|\text{Forb}_3(n, C_k)| = 2^{\Theta(n^2)}$  when  $k$  is even. We conjecture that a similar result holds for  $r > 3$  and cycles of odd length.

**Conjecture 3** *For fixed  $r \geq 3$  and  $k \geq 3$  we have  $|\text{Forb}_r(n, C_k)| = 2^{\Theta(n^{r-1})}$ .*

Almost all recent developments in this area have relied on the method of hypergraph containers that we mentioned above. What is perhaps surprising about the current work is that the proofs do not use hypergraph containers. Instead, our methods employ old and new tools in extremal (hyper)graph theory. The old tools include the extremal numbers for cycles modulo  $h$  and results about decomposing complete  $r$ -graphs into  $r$ -partite ones, and the new tools include the analysis of the shadow for extremal hypergraph problems and quantitative estimates for the bipartite canonical Ramsey problem.

## 1.1 Definitions and notations

Throughout this paper, we let  $[n]$  denote the set  $\{1, 2, \dots, n\}$ . Write  $\binom{X}{r} = \{S \subset X : |S| = r\}$  and  $\binom{X}{\leq r} = \{S \subset X : |S| \leq r\}$ . For  $X \subset [n]$ , an  $r$ -uniform hypergraph or  $r$ -graph  $H$  on vertex set  $X$  is a collection of  $r$ -element subsets of  $X$ , i.e.  $H \subset \binom{X}{r}$ . The vertex

set  $X$  is denoted by  $V(H)$ . The  $r$ -sets contained in  $H$  are *edges*. The *size* of  $H$  is  $|H|$ . Given  $S \subset V(H)$ , the *neighborhood*  $N_H(S)$  of  $S$  is the set of all  $T \subset V(H) \setminus S$  such that  $S \cup T \in H$ . The *codegree* of  $S$  is  $d_H(S) = |N_H(S)|$ . When the underlying hypergraph is clear from context, we may omit the subscripts in these definitions and write  $N(S)$  and  $d(S)$  for simplicity. The *sub-edges* of  $H$  are the  $(r-1)$ -subsets of  $[n]$  with positive codegree in  $H$ . The set of all sub-edges of  $H$  is called the *shadow* of  $H$ , and is denoted  $\partial H$ .

An  *$r$ -partite  $r$ -graph*  $H$  is an  $r$ -graph with vertex set  $\bigsqcup_{i=1}^r V_i$  (the  $V_i$ s are pairwise disjoint), and every  $e \in H$  satisfies  $|e \cap V_i| = 1$  for all  $i \in [r]$ . When all such edges  $e$  are present,  $H$  is called a *complete  $r$ -partite  $r$ -graph*. When  $|V_i| = s$  for all  $i \in [r]$ , a complete  $r$ -partite  $r$ -graph  $H$  is said to be *balanced*, and denoted  $K_{s;r}$ .

For each integer  $k \geq 1$ , a (*loose, or linear*) *path of length  $k$*  denoted by  $P_k$ , is a collection of  $k$  edges  $e_1, e_2, \dots, e_k$  such that  $|e_i \cap e_j| = 1$  if  $i = j + 1$ , and  $e_i \cap e_j = \emptyset$  otherwise.

We will often omit floors and ceilings in our calculations for ease of notation and all logs will have base 2.

## 2 Proof of the main result

The proof of Theorem 2 proceeds by counting edge-colored  $(r-1)$ -graphs with certain restrictions; the details differ quite substantially for the cases  $r = 3$  and  $r > 3$ . In this section we state the main technical statement (Theorem 5) about these edge-colorings that will be needed, as well as some other tools, and then prove Theorem 2 using these results.

### 2.1 Main technical statement

Given an  $(r-1)$ -graph  $G$  with  $V(G) \subset [n]$ , a *coloring function* is a function  $\chi : G \rightarrow [n]$  such that  $\chi(e) = z_e \in [n] \setminus e$  for every  $e \in G$ . We call  $z_e$  the *color* of  $e$ . The vector of colors  $N_G = (z_e)_{e \in G}$  is called an *edge-coloring* of  $G$ . The pair  $(G, N_G)$  is an *edge-colored  $(r-1)$ -graph*. A *color class* is the set of all edges that receive the same color.

Given  $G$ , each edge-coloring  $N_G$  defines an  $r$ -graph  $H(N_G) = \{e \cup \{z_e\} : e \in G\}$ , called the *extension of  $G$  by  $N_G$* . When there is only one coloring that has been defined, we also use the notation  $G^* = H(N_G)$  for the extension. Observe that any subgraph  $G' \subset G$  also admits an extension by  $N_G$ , namely,  $G'^* = \{e \cup \{z_e\} : e \in G'\} \subset G^*$ . If  $G' \subset G$  and  $\chi|_{G'}$  is one-to-one, then  $G'$  is called *rainbow colored*. If a rainbow colored  $G'$  further satisfies that  $z_e \notin V(G')$  for all  $e \in G'$ , then  $G'$  is said to be *strongly rainbow colored*. Note that a strongly rainbow colored graph  $C_k \subset G'$  gives rise to 3-graph  $C_k$  in  $G'^* \subset G^*$ .

**Definition 4** Given  $r \geq 3, k \geq 3, s \geq 1$ , let  $f_r(n, k, s)$  be the number of edge-colored balanced complete  $(r-1)$ -partite  $(r-1)$ -graphs  $G = K_{s;r-1}$  with  $V(G) \subset [n]$ , whose extension  $G^*$  is  $C_k$ -free.

The function  $f_r(n, k, s)$  allows us to encode  $r$ -graphs, and our main technical theorem gives an upper bound for this function.

**Theorem 5** *Given  $r \geq 3, k \geq 3$  there exist  $D = D_{k/2}$ ,  $c_2 = c_2(r, k)$ , such that*

$$f_r(n, k, s) \leq \begin{cases} 2^{(5/2)ks \log n + 4s^2 \log D} & \text{if } r = 3, k \text{ is even,} \\ 2^{(c_2+2r)s^{r-2} \log n + s^{r-1}(\log(c_2+r) + (r-2) \log s)} & \text{if } r > 3. \end{cases}$$

For  $r$  and  $k$  fixed, the bounds above can be written as

$$f_r(n, k, s) = \begin{cases} 2^{O(s \log n + s^2)} & \text{if } r = 3, k \text{ is even,} \\ 2^{O(s^{r-2} \log n + s^{r-1} \log s)} & \text{if } r > 3. \end{cases}$$

Note that the trivial upper bound is  $f_r(n, k, s) \leq n^{(r-1)s + s^{r-1}} \sim 2^{s^{r-1} \log n}$  (first choose  $(r-1)s$  vertices, then color each of its  $s^{r-1}$  edges) so Theorem 5 is nontrivial only if  $s = o(n)$ . The proof of Theorem 5 will be given in Sections 3–6.

## 2.2 Decomposing $r$ -graphs into balanced complete $r$ -partite $r$ -graphs

Chung-Erdős-Spencer [19] and Bublitz [3] proved that the complete graph  $K_n$  can be decomposed into balanced complete bipartite graphs such that the sum of the sizes of the vertex sets in these bipartite graphs is at most  $O(n^2 / \log n)$ . See also [55, 43] for some generalizations and algorithmic consequences. We need the following generalization of this result to  $r$ -graphs.

**Theorem 6** *Let  $n \geq r \geq 2$ . There exists a constant  $c'_1 = c'_1(r)$ , such that any  $n$ -vertex  $r$ -graph  $H$  can be decomposed into balanced complete  $r$ -partite  $r$ -graphs  $K_{s_i:r}, i = 1, \dots, m$ , with  $s_i \leq (\log n)^{1/(r-1)}$  and  $\sum_{i=1}^m s_i^{r-1} \leq c'_1 n^r / (\log n)^{1/(r-1)}$ .*

*Proof.* An old result of Erdős [27] states that for any integers  $r, s \geq 2$  and  $n \geq rs$ , we have

$$\text{ex}_r(n, K_{s:r}) < n^{r-1/s^{r-1}}.$$

Note that for  $r = 2$ , this was proved much earlier by Kővári-Sós-Turán [36].

We first assume that  $n \geq 2r$ . Taking the derivative, one can show that for each  $r \geq 2$ ,  $n \mapsto n/r - (\log n)^{1/(r-1)}$  is an increasing function in  $n$ , hence its minimum is achieved at  $n = 2r$ . So for all  $n \geq 2r$ , we have

$$\frac{n}{r} - (\log n)^{1/(r-1)} \geq \frac{2r}{r} - (\log 2r)^{1/(r-1)} = 2 - (\log 2r)^{1/(r-1)} \geq 0.$$

Thus, for any  $s \leq (\log n)^{1/(r-1)} \leq n/r$ , the Turán number for  $K_{s:r}$  is

$$\text{ex}_r(n, K_{s:r}) < n^{r-1/s^{r-1}}.$$

Next, we give an algorithm of decomposing  $H$  into  $K_{s:r}$ s. Let  $H_1 = H$ . For  $i \geq 1$ , repeatedly find a  $K_{s_i:r} \subset H_i$  with maximum  $s_i$  subject to  $s_i \leq (\log n)^{1/(r-1)}$  and delete it from  $H_i$  to form  $H_{i+1}$ . The loop terminates at step  $i$  if  $|H_i| \leq n^r / (\log n)^{1/(r-1)}$ . Then let the remaining graph be decomposed into single edges ( $K_{1:r}$ s).

By the algorithm, the vertex size of each  $K_{s_i:r}$  satisfies that  $s_i \leq (\log n)^{1/(r-1)}$ ,  $1 \leq i \leq m$  automatically. So we are left to show the upper bound for  $\sum_{i=1}^m s_i^{r-1}$ .

We divide the iterations of the above algorithm into phases, where the  $k$ th phase consists of those iterations where the number of edges in the input  $r$ -graph of the algorithm lies in the interval  $(n^r/(k+1), n^r/k]$ . In other words, in phase  $k$ , each  $K_{s_i:r}$  to be found is in an  $r$ -graph  $H_i$  with  $|H_i| > n^r/(k+1)$ . Define  $s(k) = (\log n / \log(k+1))^{1/(r-1)}$ . Then it is easy to see  $s(k) \leq (\log n)^{1/(r-1)} \leq n/r$ . So by Erdős' result

$$\frac{n^r}{k+1} = n^{r-1/s(k)^{r-1}} > \text{ex}(n, K_{s(k):r}).$$

Hence,  $K_{s(k):r} \subset H_i$ . So in phase  $k$ , the minimum  $s_i$  of a  $K_{s_i:r}$  we are able to find has the lower bound

$$s_i \geq s(k) = \left( \frac{\log n}{\log(k+1)} \right)^{1/(r-1)}.$$

Now notice that

$$\sum_{i=1}^m s_i^{r-1} = \sum_{i=1}^m s_i^r \cdot \frac{1}{s_i}.$$

Dividing up the terms in the summation according to phases, we observe that this is a sum of the number of edges deleted in the  $k$ th phase times a weight of  $1/s_i$  for each edge. Also notice that there are at most  $n^r/(\log n)^{1/(r-1)}$  single edges, we have

$$\begin{aligned} \sum_{i=1}^m s_i^{r-1} &\leq \frac{n^r}{(\log n)^{1/(r-1)}} + \sum_{k=1}^{\infty} n^r \left( \frac{1}{k} - \frac{1}{k+1} \right) \left( \frac{\log(k+1)}{\log n} \right)^{1/(r-1)} \\ &= \frac{n^r}{(\log n)^{1/(r-1)}} + \sum_{k=1}^{\infty} \frac{n^r (\log(k+1))^{1/(r-1)}}{k(k+1) (\log n)^{1/(r-1)}} \\ &= \frac{c'_1 n^r}{(\log n)^{1/(r-1)}}, \end{aligned}$$

where

$$c'_1 = 1 + \sum_{k=1}^{\infty} \frac{(\log(k+1))^{1/(r-1)}}{k(k+1)}.$$

Finally, for  $n < 2r$ , we just let  $H$  be decomposed into  $K_{1:r}$ s. Then we obtain the following bound for  $\sum_{i=1}^m s_i^{r-1}$ .

$$\sum_{i=1}^m s_i^{r-1} = m \leq \binom{2r}{r} \leq \frac{c'_1 n^r}{(\log n)^{1/(r-1)}}$$

with appropriately chosen  $c'_1 = c'_1(r)$ . This completes the proof.  $\square$

### 2.3 A corollary of Theorem 5

Theorem 5 is about the number of ways to edge-color complete  $(r-1)$ -partite  $(r-1)$ -graphs with parts of size  $s$  and vertex set in  $[n]$ . In this section, we use Theorems 5 and 6 to prove a related statement where we do not require the  $(r-1)$ -partite condition and the restriction to  $s$  vertices.

**Definition 7** For  $r \geq 3$  and  $k \geq 3$ , let  $g_r(n, k)$  be the number of edge-colored  $(r-1)$ -graphs  $G$  with  $V(G) \subset [n]$  such that the extension  $G^*$  is  $C_k$ -free.

**Lemma 8** Let  $r \geq 3, k \geq 3$ , and  $n$  be large enough. Then there exist  $c_1 = c_1(r)$ ,  $c_2 = c_2(r, k)$  and  $D = D_{k/2}$ , such that

$$g_r(n, k) \leq \begin{cases} 2^{(3kc_1 + 4 \log D)n^2} & \text{if } r = 3, k \text{ is even,} \\ 2^{2(c_2 + 2r)c_1 n^{r-1} (\log n)^{(r-3)/(r-2)}} & \text{if } r > 3. \end{cases}$$

Note that if  $r$  and  $k$  are fixed, for both cases we have

$$g_r(n, k) = 2^{O(n^{r-1} (\log n)^{(r-3)/(r-2)})}.$$

*Proof.* Given any  $(r-1)$ -graph  $G$ , by applying Theorem 6 with parameter  $r-1$  instead of  $r$ , we may decompose  $G$  into balanced complete  $(r-1)$ -partite  $(r-1)$ -graphs  $K_{s_1:r-1}, \dots, K_{s_m:r-1}$ , with  $s_i \leq (\log n)^{1/(r-2)}$  and  $\sum_{i=1}^m s_i^{r-2} \leq c_1 n^{r-1} / (\log n)^{1/(r-2)}$ , where  $c_1 = c_1(r) = c'_1(r-1)$ . Then we trivially deduce the following two facts.

- From the second inequality, we have  $m \leq c_1 n^{r-1} / (\log n)^{1/(r-2)}$ .
- Using the fact that these copies of  $K_{s_i:r-1}$  are edge disjoint, we have

$$\sum_{i=1}^m s_i^{r-1} \leq \binom{n}{r-1} < n^{r-1}.$$

Therefore, to construct an edge-colored  $G$ , we need to first choose a sequence of positive integers  $(m, s_1, \dots, s_m)$  such that  $m \leq c_1 n^{r-1} / (\log n)^{1/(r-2)}$ , and  $s_i \leq (\log n)^{1/(r-2)}$  for all  $i$ . More formally, let

$$S_{n,r} = \{(m, s_1, s_2, \dots, s_m) : m \leq c_1 n^{r-1} / (\log n)^{1/(r-2)}, 1 \leq s_i \leq (\log n)^{1/(r-2)}, 1 \leq i \leq m\}.$$

Then

$$\begin{aligned} |S_{n,r}| &\leq \frac{c_1 n^{r-1}}{(\log n)^{1/(r-2)}} \left( (\log n)^{1/(r-2)} \right)^{\frac{c_1 n^{r-1}}{(\log n)^{1/(r-2)}}} \\ &= 2^{\log \left( \frac{c_1 n^{r-1}}{(\log n)^{1/(r-2)}} \right) + \frac{c_1 n^{r-1} \log(\log n)^{1/(r-2)}}{(\log n)^{1/(r-2)}}} \\ &\leq 2^{c_1 n^{r-1}}. \end{aligned}$$

Then, we sequentially construct edge-colored  $K_{s_i:r-1}$  for each  $i \in [m]$ . To make sure  $G^*$  is  $C_k$ -free,  $K_{s_i:r-1}^*$  has to be made  $C_k$ -free in the first place. Applying Theorem 5, we get the following upper bounds. For  $r = 3$  and even  $k$ ,

$$\begin{aligned}
g_3(n, k) &\leq \sum_{(m, s_1, \dots, s_m) \in S_{n,3}} \prod_{i=1}^m f_3(n, k, s_i) \\
&\leq \sum_{(m, s_1, \dots, s_m) \in S_{n,3}} \prod_{i=1}^m 2^{(5/2)ks_i \log n + 4s_i^2 \log D} \\
&\leq \sum_{(m, s_1, \dots, s_m) \in S_{n,3}} 2^{\sum_{i=1}^m (5/2)ks_i \log n + 4s_i^2 \log D} \\
&\leq \sum_{(m, s_1, \dots, s_m) \in S_{n,3}} 2^{(5/2)k \log n (c_1 n^2 / \log n) + 4n^2 \log D} \\
&\leq 2^{c_1 n^2} \cdot 2^{(5/2)kc_1 n^2 + 4n^2 \log D} \\
&\leq 2^{(3kc_1 + 4 \log D)n^2}.
\end{aligned}$$

For  $r > 3$ , and  $(m, s_1, \dots, s_m) \in S_{n,r}$ , the number of ways to construct these copies of  $K_{s_i:r-1}$  is at most

$$\begin{aligned}
\prod_{i=1}^m f_r(n, k, s_i) &\leq \prod_{i=1}^m 2^{(c_2+2r)s_i^{r-2} \log n + s_i^{r-1} (\log(c_2+r-1) + (r-2) \log s_i)} \\
&= 2^{\sum_{i=1}^m (c_2+2r)s_i^{r-2} \log n + s_i^{r-1} (\log(c_2+r-1) + (r-2) \log s_i)} \\
&\leq 2^{(c_2+2r) \frac{c_1 n^{r-1}}{(\log n)^{1/(r-2)}} \log n + n^{r-1} (\log(c_2+r-1) + (r-2) \log(\log n)^{1/(r-2)})} \\
&\leq 2^{(c_2+2r)c_1 n^{r-1} (\log n)^{(r-3)/(r-2)} + n^{r-1} (\log(c_2+r-1) + (r-2) \log(\log n)^{1/(r-2)})} \\
&\leq 2^{(3/2)(c_2+2r)c_1 n^{r-1} (\log n)^{(r-3)/(r-2)}}.
\end{aligned}$$

Note that this is the only place in the proof where we use  $s_i \leq (\log n)^{1/(r-2)}$ .

Therefore,

$$\begin{aligned}
g_r(n, k) &\leq \sum_{(m, s_1, \dots, s_m) \in S_{n,r}} \prod_{i=1}^m f_r(n, k, s_i) \\
&\leq 2^{c_1 n^{r-1}} \cdot 2^{(3/2)(c_2+2r)c_1 n^{r-1} (\log n)^{(r-3)/(r-2)}} \\
&\leq 2^{2(c_2+2r)c_1 n^{r-1} (\log n)^{(r-3)/(r-2)}},
\end{aligned}$$

and the proof is complete.  $\square$

## 2.4 Finding a cycle if codegrees are high

A crucial statement that we use in our proof is that any  $r$ -graph such that every sub-edge has high codegree contains rich structures, including cycles. This was explicitly proved in [37] and we reproduce the proof here for completeness.



**Lemma 9 (Lemma 3.2 in [37])** *For  $r, k \geq 3$ , if all sub-edges of an  $r$ -graph  $H$  have codegree greater than  $rk$ , then  $C_k \subset H$ .*

*Proof.* Let  $F = \partial^{r-2}H$  be the  $(2)$ -graph that consists of pairs that are contained in some edge of  $H$ . Note that each edge of  $H$  induces a  $K_r$  in  $F$ , so all edges of  $F$  are contained in some triangle ( $C_3$ ). Furthermore, since all sub-edges of  $H$  have codegree greater than  $rk$ , each edge of  $F$  is in more than  $rk$  triangles. We will first find a  $k$ -cycle in  $F$  as follows. Starting with a triangle  $C_3$ , for  $i = 3, \dots, k-1$  pick an edge  $e \in C_i$ , form  $C_{i+1}$  by replacing  $e$  by the other two edges of one of at least  $rk - i + 2$  triangles containing  $e$  and excluding other vertices of  $C_i$ .

Next, let  $C_k \subset F$  be a  $k$ -cycle with edges  $f_1, \dots, f_k$ . Find in  $H$  a subgraph  $C = \{e_i : e_i = f_i \cup g_i, i \in [k]\}$  such that  $V(C) = \cup_{i=1}^k e_i$  is of maximum size. Suppose  $C$  is not a  $k$ -cycle in  $H$ . Then there are distinct  $i, j$  such that  $g_i \cap g_j \neq \emptyset$ . Pick  $v \in g_i \cap g_j$  and consider the sub-edge  $e_i \setminus \{v\} = f_i \cup g_i \setminus \{v\}$ . The codegree  $d_H(e_i \setminus \{v\}) > rk$  by assumption. On the other hand,  $|V(C)| < rk$  since  $C$  is not a  $k$ -cycle, so there exists a vertex  $v' \in N_H(e_i \setminus \{v\}) \setminus V(C)$ . Replacing  $e_i$  by  $e_i \setminus \{v\} \cup \{v'\}$ , we obtain a  $C'$  with a larger vertex set, a contradiction. So  $H$  contains a  $C_k$ .  $\square$

## 2.5 Proof of Theorem 2

Now we have all the ingredients to complete the proof of our main result.

**Proof of Theorem 2.** Starting with any  $r$ -graph  $H$  on  $[n]$  with  $C_k \not\subset H$ , we claim that there exists a sub-edge with codegree at most  $rk$ . Indeed, otherwise all sub-edges of  $H$  will have codegree more than  $rk$ , and then by Lemma 9 we obtain a  $C_k \subset H$ . Let  $e'$  be the sub-edge of  $H$  with  $0 < d_H(e') \leq rk$  such that it has smallest lexicographic order among all such sub-edges. Delete all edges of  $H$  containing  $e'$  from  $H$  (i.e. delete  $\{e \in H : e' \subset e\}$ ). Repeat this process of “searching and deleting” in the remaining  $r$ -graph until there are no such sub-edges. We claim that the remaining  $r$ -graph must have no edges at all. Indeed, otherwise we get a nonempty subgraph all of whose sub-edges have codegree greater than  $rk$ , and again by Lemma 9, we obtain a  $C_k \subset H$ .

Given any  $C_k$ -free  $r$ -graph  $H$  on  $[n]$ , the algorithm above sequentially decomposes  $H$  into a collection of sets of at most  $rk$  edges who share a sub-edge (an  $(r-1)$ -set) in common. We regard the collection of these  $(r-1)$ -sets as an  $(r-1)$ -graph  $G$ . Moreover, for each edge  $e \in G$ , let  $N_e$  be the set of vertices  $v \in V(H)$  such that  $e \cup \{v\}$  is an edge of  $H$  at the time  $e$  was chosen. So  $N_e \in \binom{[n] \setminus e}{\leq rk}$ , for all  $e \in G$ . Thus, we get a map

$$\phi : \text{Forb}_r(n, C_k) \longrightarrow \left\{ (G, N_G) : G \subset \binom{[n]}{r-1}, N_G = \left( N_e \in \binom{[n] \setminus e}{\leq rk} : e \in G \right) \right\}.$$

We observe that  $\phi$  is injective. Indeed,

$$\phi^{-1}((G, N_G)) = H(N_G) = \{e \cup \{z_e\} : e \in G, z_e \in N_e\},$$

therefore  $|\text{Forb}_r(n, C_k)| = |\phi(\text{Forb}_r(n, C_k))|$ . Let  $P = \phi(\text{Forb}_r(n, C_k))$  which is the set of all pairs  $(G, N_G)$  such that  $H(N_G)$  is  $C_k$ -free. Next we describe our strategy for upper bounding  $|P|$ .

For each pair  $(G, N_G) \in P$  and  $e \in G$ , we pick exactly one  $z_e^1 \in N_e$ . Thus we get a pair  $(G_1, N_{G_1})$ , where  $G_1 = G$ , and  $N_{G_1} = (z_e^1 : e \in G_1)$ . Then, delete  $z_e^1$  from each  $N_e$ , let  $G_2 = \{e \in G_1 : N_e \setminus \{z_e^1\} \neq \emptyset\}$  and pick  $z_e^2 \in N_e \setminus \{z_e^1\}$  to get the pair  $(G_2, N_{G_2})$ . For  $2 \leq i < rk$ , we repeat this process for  $G_i$  to obtain  $G_{i+1}$ . Since each  $N_{G_i}$  contains only singletons, the pair  $(G_i, N_{G_i})$  can be regarded as an edge-colored  $(r-1)$ -graph. Note that we may get some empty  $G_i$ s. This gives us a map

$$\psi : P \longrightarrow \left\{ (G_1, \dots, G_{rk}) : G_i \subset \binom{[n]}{r-1} \text{ is edge-colored for all } i \in [rk] \right\}.$$

Moreover, it is almost trivial to observe that  $\psi$  is injective, since if  $y \neq y'$ , then either the underlying  $(r-1)$ -graphs of  $y$  and  $y'$  differ, or the  $(r-1)$ -graphs are the same but the color sets differ. In both cases one can easily see that  $\psi(y) \neq \psi(y')$ . Again, we let  $Q = \psi(P)$ .

By Lemma 8, we have

$$\begin{aligned} |\text{Forb}_r(n, C_k)| = |P| = |Q| &\leq \prod_{i=1}^{rk} g_r(n, k) \\ &\leq \begin{cases} \prod_{i=1}^{3k} 2^{(3kc_1 + 4 \log D)n^2} & \text{if } r = 3, k \text{ is even} \\ \prod_{i=1}^{rk} 2^{2(c_2 + 2r)c_1 n^{r-1} (\log n)^{(r-3)/(r-2)}} & \text{if } r > 3 \end{cases} \\ &\leq \begin{cases} 2^{\sum_{i=1}^{3k} (3kc_1 + 4 \log D)n^2} & \text{if } r = 3, k \text{ is even} \\ 2^{\sum_{i=1}^{rk} 2(c_2 + 2r)c_1 n^{r-1} (\log n)^{(r-3)/(r-2)}} & \text{if } r > 3 \end{cases} \\ &\leq \begin{cases} 2^{(9k^2 c_1 + 12k \log D)n^2} & \text{if } r = 3, k \text{ is even} \\ 2^{2rk(c_2 + 2r)c_1 n^{r-1} (\log n)^{(r-3)/(r-2)}} & \text{if } r > 3 \end{cases} \\ &= 2^{c n^{r-1} (\log n)^{(r-3)/(r-2)}}. \end{aligned}$$

where  $c = 9k^2 c_1 + 12k \log D$  if  $r = 3$  and  $k$  is even, and  $c = 2rk(c_2 + 2r)c_1$  if  $r > 3$ . The constants  $c_1 = c_1(r)$ ,  $c_2 = c_2(r, k)$  and  $D = D_{k/2}$  are from Theorem 6 and Theorem 5, respectively.  $\square$

### 3 Proof of Theorem 5 for $r > 3$

In the remaining part of the paper, we prove Theorem 5. The cases  $r = 3$  and  $r > 3$  have quite different proofs. In this short section we prove the case  $r > 3$ .

We need one more ingredient to prove Theorem 5, namely a partite version of the extremal result for  $C_{2l}$ . This is a corollary of the main result of [37] although it can also be proved directly by analyzing the shadow with a much better bound.

**Lemma 10** *Let  $r \geq 3$ ,  $k \geq 3$  and  $G$  be an  $r$ -partite  $r$ -graph on vertex sets  $\bigsqcup_{i=1}^r V_i$  with  $|V_i| = s$ , for all  $i$ . There exists  $c'_2 = c'_2(r, k)$ , such that if  $|G| > c'_2 s^{r-1}$  then  $G$  contains a cycle of length  $k$ .*

*Proof.* By Theorem 1.1 of [37], we have

$$\text{ex}_r(n, C_k) \sim \binom{n}{r} - \binom{n - \lfloor (k-1)/2 \rfloor}{r} \sim \frac{\lfloor (k-1)/2 \rfloor}{(r-1)!} n^{r-1}.$$

So, we may take  $c'_2$  large enough such that

$$c'_2 s^{r-1} = \frac{c'_2}{r^{r-1}} n^{r-1} \geq \frac{\lfloor (k-1)/2 \rfloor}{(r-1)!} n^{r-1},$$

to guarantee the existence of a copy of  $C_k$  in  $G$ . We remark that  $c'_2 = \frac{\lfloor k/2 \rfloor r^{r-1}}{(r-1)!}$  suffices.  $\square$

**Proof of Theorem 5 for  $r > 3$ .** Let  $G = K_{s:r-1}$  with  $V(G) \subset [n]$  and  $C_k \not\subset G^*$ . For any edge-coloring  $N_G = (z_e : e \in G)$  of  $G$ , let  $Z = \{z_e : e \in G\} \subset [n]$  be the set of all its colors. We first argue that  $|Z| < (c_2 + r - 1)s^{r-2}$ , where  $c_2 = c_2(r, k) = c'_2(r-1, k) = \lfloor k/2 \rfloor (r-1)^{r-2}/(r-2)!$ , the constant from Lemma 10. Indeed, if  $|Z| \geq (c_2 + r - 1)s^{r-2}$ , then  $|Z \setminus V(G)| \geq (c_2 + r - 1)s^{r-2} - s(r-1) > (c_2 + r - 1 - (r-1))s^{r-2} = c_2 s^{r-2}$ . For each color  $v \in Z \setminus V(G)$  pick an edge in  $G$  with color  $v$ . We get a subgraph  $G' \subset G$  that is strongly rainbow with  $|G'| = |Z \setminus V(G)| > c_2 s^{r-2}$ . By Lemma 10, we find an  $(r-1)$ -uniform  $C_k$  in  $G'$  that is strongly rainbow, which contradicts the fact that  $C_k \not\subset G^*$ .

We now count the number of edge-colored  $K_{s:r-1}$  as follows: first choose  $s(r-1)$  vertices from  $[n]$  as the vertex set, then choose at most  $(c_2 + r - 1)s^{r-2}$  colors, finally color each edge of the  $K_{s:r-1}$ . As  $|K_{s:r-1}| = s^{r-1}$ , this yields

$$\begin{aligned} f_r(n, k, s) &\leq n^{s(r-1) + (c_2 + r - 1)s^{r-2}} ((c_2 + r - 1)s^{r-2})^{s^{r-1}} \\ &= 2^{(s(r-1) + (c_2 + r - 1)s^{r-2}) \log n + s^{r-1} (\log(c_2 + r - 1) + (r-2) \log s)} \\ &\leq 2^{(c_2 + 2r)s^{r-2} \log n + s^{r-1} (\log(c_2 + r - 1) + (r-2) \log s)}. \end{aligned} \quad \square$$

## 4 Proof of Theorem 5 for $r = 3$ and even $k$

The rest of the paper is devoted to the proof of Theorem 5 for  $r = 3$  and even  $k$ . For simplicity of presentation, we write  $k = 2l$  where  $l \geq 2$ . Our two main tools are the following lemmas about edge-coloring bipartite graphs.

**Lemma 11** *Let  $l \geq 2, s, t \geq 1$ ,  $G = K_{s,t}$  be an edge-colored complete bipartite graph with  $V(G) \subset [n]$  and  $Z = \{z_e : e \in G\} \subset [n]$  be the set of all colors. If  $G$  contains no strongly rainbow colored  $C_{2l}$ , then  $|Z| < 2l(s+t)$ .*

**Lemma 12** *For each  $l \geq 2$ , there exists a constant  $D = D_l > 0$ , such that the following holds. Let  $s, t \geq 1$ ,  $G = K_{s,t}$  be a complete bipartite graph with vertex set  $V(G) \subset [n]$ , and  $Z \subset [n]$  be a set of colors. Then the number of ways to edge-color  $G$  with  $Z$  such that the extension  $G^*$  contains no  $C_{2l}$ , is at most  $D^{(s+t)^2}$ .*

The proofs of these lemmas require several new ideas which will be presented in the rest of the paper. Here we quickly show that they imply Theorem 5 for  $r = 3$  and even  $k$ .

**Proof of Theorem 5 for  $r = 3$  and  $k = 2l$ .** Recall that  $r \geq 3$ ,  $l \geq 2$ , and that  $f_3(n, 2l, s)$  is the number of edge-colored copies of  $K_{s,s}$  whose vertex set lies in  $[n]$  and whose (3-uniform) extension is  $C_{2l}$ -free. To obtain such a copy of  $K_{s,s}$ , we first choose from  $[n]$  its  $2s$  vertices, then its at most  $4ls$  colors by Lemma 11 and finally we color this  $K_{s,s}$  by Lemma 12. This yields

$$f_3(n, 2l, s) \leq n^{2s+4ls} D^{(2s)^2} \leq 2^{5ls \log n + 4s^2 \log D} = 2^{(5/2)ks \log n + 4s^2 \log D},$$

where the last inequality holds since  $l \geq 2$ . Note that  $D = D_l = D_{k/2}$  is the desired constant.  $\square$

## 5 Proof of Lemma 11

In this section we prove Lemma 11. Our main tool is an extremal result about cycles modulo  $h$  in a graph. This problem has a long history, beginning with a Conjecture of Burr and Erdős that was solved by Bollobás [2] in 1976 via the following result: for each integer  $m$  and odd positive integer  $h$ , every graph  $G$  with minimum degree  $\delta(G) \geq 2((h+1)^h - 1)/h$  contains a cycle of length congruent to  $m$  modulo  $h$ . The lower bound on  $\delta(G)$  was improved by Thomassen [53] who also generalized it to the case with all integers  $h$ . It was conjectured by Thomassen that graphs with minimum degree at least  $h+1$  contain a cycle of length  $2m$  modulo  $h$ , for any  $h, m \geq 1$ . The conjecture received new attention recently. In particular Liu-Ma [41] settled the case when  $h$  is even, and Diwan [24] proved it for  $m = 2$ . To date, Sudakov and Verstraëte [52] hold the best known bound for the general case on this problem.

We need the very special case  $m = 1$  of Thomassen's conjecture and in order to be self contained, we give a proof below. The idea behind this proof can be found in Diwan [24].

**Lemma 13** *If  $G$  is an  $n$ -vertex graph with at least  $(h+1)n$  edges, then  $G$  contains a cycle of length 2 modulo  $h$ .*

*Proof.* By removing vertices of degree at most  $h$ , we may assume that  $G$  has minimum degree at least  $h+1$ . Let  $P$  be a longest path in  $G$ . Assume that  $P$  is of length  $l$ . Let  $V(P) = \{x_0, x_1, \dots, x_l\}$ , where  $x_0, x_l$  are the two end-vertices of  $P$ , and  $x_{i-1}x_i \in P$  for all  $i \in [l]$ . Then we observe that  $N(x_0) \subset V(P)$ . Otherwise we can extend  $P$  to a longer path by  $x_0y$  with some vertex  $y \in N(x_0) \setminus V(P)$ . So  $N(x_0) = \{x_i : i \in I\}$  for some  $I \subset [l]$ . Note that  $1 \in I$ ,  $|I| = |N(x_0)| \geq \delta(G) \geq h+1$ , and the distance  $\text{dist}_P(x_i, x_j) = |j - i|$ . Consider the set  $J = \{i - 1 : i \in I, i \neq 1\}$ , note that this is the set of all distances  $\text{dist}_P(x_1, x_i)$  with  $x_i \in N(x_0)$  and  $i \neq 1$ . Clearly,  $|J| = |I| - 1 \geq h$ . If there exists  $i - 1 \in J$  with  $i - 1 \equiv 0 \pmod{h}$ , then we are done, since the sub-path of  $P$  from  $x_1$  to  $x_i$  together with  $x_0x_1, x_0x_i$  form a cycle of length 2 modulo  $h$ . So none of the numbers in  $J$  are multiples of  $h$ . By the pigeonhole principle, there are at least two elements  $i - 1, j - 1 \in J$  such that  $i - 1 \equiv j - 1 \pmod{h}$ , thus  $\text{dist}_P(x_i, x_j) = |j - i| \equiv 0 \pmod{h}$ . Again, we can find a cycle of length 2 modulo  $h$  by taking the sub-path of  $P$  connecting  $x_i, x_j$  and edges  $x_0x_i, x_0x_j$ .  $\square$

**Lemma 14** *Let integers  $l \geq 2$ ,  $s, t \geq 1$ ,  $G = K_{s,t}$  with  $V(G) \subset [n]$  be edge-colored. If  $G$  contains a strongly rainbow colored cycle of length  $2 \pmod{2l-2}$ , then  $G$  contains a strongly rainbow colored  $C_{2l}$ .*

*Proof.* Let us assume that  $C$  is the shortest strongly rainbow colored cycle of length 2 modulo  $2l-2$  in  $G$ . Then  $C$  has at least  $2l$  edges. We claim that  $C$  is a  $C_{2l}$ . Suppose not, let  $e$  be a chord of  $C$  (such a chord exists as  $G$  is complete bipartite), such that  $C$  is cut up into two paths  $P_1$  and  $P_2$  by the two endpoints of  $e$ , and  $|P_1| = 2l-1$ . Let  $Z_1, Z_2$  be the set of their colors respectively. If the color  $z_e \notin Z_1 \cup V(P_1) \setminus e$ , then  $P_1 \cup e$  is a strongly rainbow colored cycle of length  $2l$ , a contradiction. Therefore  $z_e \in Z_1 \cup V(P_1) \setminus e$ , but then  $z_e \notin Z_2 \cup V(P_2) \setminus e$ , yielding a shorter strongly rainbow colored cycle  $P_2 \cup e$  of length 2 modulo  $2l-2$ , a contradiction.  $\square$

We now have all the necessary ingredients to prove Lemma 11.

**Proof of Lemma 11.** Suppose that  $|Z| \geq 2l(s+t)$ . Then  $|Z \setminus V(G)| \geq (2l-1)(s+t)$ . For each color  $v$  in  $Z \setminus V(G)$ , pick an edge  $e$  of  $G$  with color  $v$ . We obtain a strongly rainbow colored subgraph  $G'$  of  $G$  with at least  $(2l-1)(s+t)$  edges. Lemma 13 guarantees the existence of a rainbow colored cycle of length 2 modulo  $2l-2$  in  $G'$ . By construction, this cycle is strongly rainbow. Lemma 14 then implies that there is a strongly rainbow colored  $C_{2l}$  in  $G$ .  $\square$

## 6 Proof of Lemma 12

Our proof of Lemma 12 is inspired by the methods developed in [37]. The main idea is to use the bipartite canonical Ramsey theorem. In order to use this approach we need to develop some new quantitative estimates for an asymmetric version of the bipartite canonical Ramsey theorem.

### 6.1 Canonical Ramsey theory

In this section we state and prove the main result in Ramsey theory that we will use to prove Lemma 12. We are interested in counting the number of edge-colorings of a bipartite graph, such that the (3-uniform) extension contains no copy of  $C_{2l}$ . The canonical Ramsey theorem allows us to find nice colored structures that are easier to work with. However, the quantitative aspects are important for our application and consequently we need to prove various bounds for bipartite canonical Ramsey numbers. We begin with some definitions.

Let  $G$  be a bipartite graph on vertex set with bipartition  $X \sqcup Y$ . For any subsets  $X' \subset X$ ,  $Y' \subset Y$ , let  $E_G(X', Y') = G[X' \sqcup Y'] = \{xy \in G : x \in X', y \in Y'\}$ , and  $e_G(X', Y') = |E_G(X', Y')|$ . If  $X'$  contains a single vertex  $x$ , then  $E_G(\{x\}, Y')$  will be simply written as  $E_G(x, Y')$ . The subscript  $G$  may be omitted if it is obvious from context.

**Definition 15** *Let  $G$  be an edge-colored bipartite graph with  $V(G) = X \sqcup Y$ .*

- $G$  is monochromatic if all edges in  $E(X, Y)$  are colored by the same color.
- $G$  is weakly  $X$ -canonical if  $E(x, Y)$  is monochromatic for each  $x \in X$ .
- $G$  is  $X$ -canonical if it is weakly  $X$ -canonical and for all distinct  $x, x' \in X$  the colors used on  $E(x, Y)$  and  $E(x', Y)$  are all different.

In all these cases, the color  $z_x$  of the edges in  $E(x, Y)$  is called a canonical color.

**Lemma 16** *Let  $G = K_{a,b}$  be an edge-colored complete bipartite graph with bipartition  $A \sqcup B$ , with  $|A| = a, |B| = b$ . If  $G$  is weakly  $A$ -canonical, then there exists a subset  $A' \subset A$  with  $|A'| = \sqrt{a}$  such that  $G[A' \sqcup B] = K_{\sqrt{a},b}$  is  $A'$ -canonical or monochromatic.*

*Proof.* Take a maximal subset  $A'$  of  $A$  such that the coloring on  $E(A', B)$  is  $A'$ -canonical. If  $|A'| \geq \sqrt{a}$ , then we are done. So, we may assume that  $|A'| < \sqrt{a}$ . By maximality of  $A'$ , there are less than  $\sqrt{a}$  canonical colors. By the pigeonhole principle, there are at least  $|A|/|A'| \geq a/\sqrt{a} = \sqrt{a}$  vertices of  $A$  sharing the same canonical color, which gives a monochromatic  $K_{\sqrt{a},b}$ .  $\square$

Our next lemma guarantees that in an “almost” rainbow colored complete bipartite graph, there exists a rainbow complete bipartite graph.

**Lemma 17** *For any integer  $c \geq 2$ , and  $p > c^4$ , if  $G = K_{p,p}$  is an edge-colored complete bipartite graph, in which each color class is a matching, then  $G$  contains a rainbow colored  $K_{c,c}$ .*

*Proof.* Let  $A \sqcup B$  be the vertex set of  $G$ . Pick two  $c$ -sets  $X, Y$  from  $A$  and  $B$  respectively at random with uniform probability. For any pair of monochromatic edges  $e, e'$ , the probability that they both appear in the induced subgraph  $E(X, Y)$  is

$$\left( \frac{\binom{p-2}{c-2}}{\binom{p}{c}} \right)^2 = \left( \frac{c(c-1)}{p(p-1)} \right)^2.$$

On the other hand, the total number of pairs of monochromatic edges is at most  $p^3/2$ , since every color class is a matching. Therefore the union bound shows that, when  $p > c^4$ , the probability that there exists a monochromatic pair of edges in  $E(X, Y)$  is at most

$$\frac{p^3}{2} \left( \frac{c(c-1)}{p(p-1)} \right)^2 = \frac{pc^4}{2(p-1)^2} < 1.$$

Consequently, there exists a choice of  $X$  and  $Y$  such that the  $E(X, Y)$  contains no pair of monochromatic edges. Such an  $E(X, Y)$  is a rainbow colored  $K_{c,c}$ .  $\square$

Now we are ready to prove the main result of this section which is a quantitative version of a result from [38]. Note that the edge-coloring in this result uses an arbitrary set of colors. Since the conclusion is about “rainbow” instead of “strongly rainbow”, it is not essential to have the set of colors disjoint from the vertex set of the graph.

**Theorem 18** (Asymmetric bipartite canonical Ramsey theorem) *For any integer  $l \geq 2$ , there exists real numbers  $\epsilon = \epsilon(l) > 0, s_0 = s_0(l) > 0$ , such that if  $G = K_{s,t}$  is an edge-colored complete bipartite graph on vertex set  $X \sqcup Y$  with  $|X| = s, |Y| = t$  with  $s > s_0$  and  $s/\log s < t \leq s$ , then one of the following holds:*

- $G$  contains a rainbow colored  $K_{4l,4l}$ ,
- $G$  contains a  $K_{q,2l}$  on vertex set  $Q \sqcup R$ , with  $|Q| = q, |R| = 2l$  that is  $Q$ -canonical, where  $q = s^\epsilon$ ,
- $G$  contains a monochromatic  $K_{q,2l}$  on vertex set  $Q \sqcup R$ , with  $|Q| = q, |R| = 2l$ , where  $q = s^\epsilon$ .

Note that in the last two cases, it could be  $Q \subset X, R \subset Y$  or the other way around.

*Proof.* We will show that  $\epsilon = 1/18l$ . First, fix a subset  $Y'$  of  $Y$  with  $|Y'| = t^{1/4l}$  and let

$$W = \left\{ x \in X : \text{there exists a } Y'' \in \binom{Y'}{2l} \text{ such that } E_G(x, Y'') \text{ is monochromatic} \right\}.$$

If  $|W| > s/2l$ , then the number of  $Y'' \in \binom{Y'}{2l}$  such that  $E_G(x, Y'')$  is monochromatic for some  $x$  (with repetition) is greater than  $s/2l$ . On the other hand,  $|\binom{Y'}{2l}| < |Y'|^{2l} = \sqrt{t}$ . By the pigeonhole principle, there exists a  $Y'' \in \binom{Y'}{2l}$  such that at least

$$\frac{s}{2l\sqrt{t}} \geq \frac{s}{2l\sqrt{s}} \geq s^{1/3}$$

vertices  $x$  have the property that  $E_G(x, Y'')$  is monochromatic. Let  $Q_1$  be a set of  $s^{1/3}$  such  $x$ . Then we obtain a weakly  $Q_1$ -canonical  $K_{s^{1/3}, 2l}$  on  $Q_1 \sqcup Y''$  which, by Lemma 16, contains a canonical or monochromatic  $K_{s^{1/6}, 2l}$ . Since  $\epsilon < 1/6$ , this contains a  $K_{s^\epsilon, 2l}$  as desired.

We may now assume that  $|W| \leq s/2l$ . By definition of  $W$  and the pigeonhole principle,  $E_G(x, Y')$  contains at least  $|Y'|/2l$  (distinct) colors for every  $x \in X \setminus W$ . Hence, for each  $x \in X \setminus W$  we can take  $|Y'|/2l$  distinctly colored edges from  $E(x, Y')$  to obtain a subgraph  $G'$  of  $G$  on  $(X \setminus W) \sqcup Y'$  with  $|X \setminus W||Y'|/2l$  edges.

Pick a subset  $X' \subset X \setminus W$  with  $|X'| = s^{1/16l^2}$  and  $e_{G'}(X', Y') \geq |X'||Y'|/2l$ . This is possible by an easy averaging argument. Let

$$Z = \left\{ y \in Y' : \text{there exists an } X'' \in \binom{X'}{2l} \text{ such that } E_{G'}(X'', y) \text{ is monochromatic} \right\}.$$

If  $|Z| > |Y'|/20l$ , then the number of  $X'' \in \binom{X'}{2l}$  such that  $E_{G'}(X'', y)$  is monochromatic for some  $y$  (with repetition) is greater than  $|Y'|/20l$ . On the other hand,  $|\binom{X'}{2l}| < |X'|^{2l} = s^{1/8l}$ . By the pigeonhole principle, there exists a  $X'' \in \binom{X'}{2l}$  such that at least

$$\frac{|Y'|}{20ls^{1/8l}} = \frac{t^{1/4l}}{20ls^{1/8l}} \geq \frac{s^{1/4l}}{(\log s)^{1/4l} 20ls^{1/8l}} \geq s^{1/9l} = s^{2\epsilon}$$



vertices  $y$  have the property that  $E_{G'}(X'', y)$  is monochromatic. Let  $Q_2$  be a set of  $s^{2\epsilon}$  such  $y$ . We find a weakly  $Q_2$ -canonical  $K_{2l, s^{2\epsilon}}$  on  $X'' \sqcup Q_2$ . Again, by Lemma 16, a copy of  $K_{2l, s^\epsilon}$  that is monochromatic or canonical is obtained.

Finally, we may assume that  $|Z| \leq |Y'|/20l$ . Then

$$\begin{aligned} e_{G'}(X', Y' \setminus Z) &\geq e_{G'}(X', Y') - |X'| |Z| \geq \frac{1}{2l} |X'| |Y'| - \frac{1}{20l} |X'| |Y'| = \frac{9}{20l} |X'| |Y'| \\ &\geq \frac{9}{20l} |X'| |Y' \setminus Z|. \end{aligned}$$

Since each vertex  $y \in Y' \setminus Z$  has the property that  $E_{G'}(X', y)$  sees each color at most  $2l - 1$  times, for each  $y \in Y' \setminus Z$  we may remove all edges from  $E_{G'}(X', y)$  with duplicated colors (keep one for each color). We end up getting a bipartite graph  $G''$  on  $X' \sqcup (Y' \setminus Z)$  with at least  $9|X'| |Y' \setminus Z|/40l^2$  edges. By the Kővári-Sós-Turán theorem [36], there is a  $c > 0$  such that  $G''$  contains a copy  $K$  of  $K_{p,p}$  where  $p > c \log s$ . Let  $V(K) = A \sqcup B$ . For each  $x \in A$ , the edges set  $E(x, B)$  is rainbow colored, and for each  $y \in B$ , the edge set  $E(A, y)$  is rainbow colored. Therefore each color class in  $K$  is a matching. By Lemma 17 and  $s > s_0 > 2^{(4l)^4/c}$ , we can find a rainbow colored  $K_{4l, 4l}$  in  $K$  as desired.  $\square$

## 6.2 The Induction argument for Lemma 12

We are now ready to prove Lemma 12. Let us recall the statement.

**Lemma 12** *For each  $l \geq 2$ , there exists a constant  $D = D_l > 0$ , such that the following holds. Let  $s, t \geq 1$ ,  $G = K_{s,t}$  be a complete bipartite graph with vertex set  $V(G) \subset [n]$ , and  $Z \subset [n]$  be a set of colors. Then the number of ways to edge-color  $G$  with  $Z$  such that the extension  $G^*$  contains no  $C_{2l}$ , is at most  $D^{(s+t)^2}$ .*

**Proof of Lemma 12.** Let the vertex set of  $G$  be  $S \sqcup T$  with  $|S| = s$  and  $|T| = t$ . We apply induction on  $s + t$ . By Lemma 11,  $|Z| := \sigma < 2l(s + t)$ . The number of ways to color  $G$  is at most  $\sigma^{st}$ . As long as  $s + t \leq D/2l$ , we have

$$\sigma^{st} \leq D^{st} \leq D^{(s+t)^2}$$

and this concludes the base case(s).

For the induction step, we may henceforth assume  $s + t > D/2l$ , and the statement holds for all smaller values of  $s + t$ . Let us also assume without loss of generality that  $t \leq s$ .

Next, we deal with the case  $t \leq s/\log s$ . Let  $D > 16l^2$ . Then  $s > (s + t)/2 > D/4l > 4l$  and the number of ways to color  $G$  is at most

$$\sigma^{st} \leq (2l(s + t))^{st} \leq 2^{\frac{s^2 \log(2l(s+t))}{\log s}} \leq 2^{\frac{s^2 \log(4ls)}{\log s}} \leq 2^{2s^2} \leq 2^{(s+t)^2 \log D} = D^{(s+t)^2}.$$

Therefore, we may assume that  $s/\log s < t \leq s$ , and  $s > D/4l > s_0(l)$  so the conditions of Theorem 18 hold. Let  $N_G = (z_e)_{e \in G}$  be an edge-coloring of  $G$  using colors in  $Z$ . By Theorem 18, such an edge-colored  $G$  will contain a subgraph  $G'$  that is either



- a rainbow colored  $K_{4l,4l}$ , or
- a  $Q$ -canonical  $K_{q,2l}$ , or
- a monochromatic  $K_{q,2l}$ ,

where  $|Q| = q = s^\epsilon$ .

**Claim 19**  $G'$  cannot be a rainbow colored  $K_{4l,4l}$ .

*Proof of Claim 19.* Suppose for a contradiction that  $G' = K_{4l,4l}$  is rainbow colored and  $Z'$  is the set of colors used on  $G'$ . Then  $|Z' \setminus V(G')| \geq 16l^2 - 8l$ . Pick an edge of each color in  $Z' \setminus V(G')$  to obtain a strongly rainbow colored subgraph  $G''$  of  $G'$  with  $|G''| = 16l^2 - 8l \geq (2l-1)8l$ . By Lemma 13,  $G''$  contains a strongly rainbow colored cycle of length  $2 \bmod 2l - 2$ . Lemma 14 now implies the existence of a strongly rainbow colored  $C_{2l}$  in  $G''$ , which forms a linear  $C_{2l}$  in  $G^*$ , a contradiction.  $\square$

Let  $q = s^\epsilon$  and  $\alpha$  be the number of edge-colorings of  $G$  that contain a  $Q$ -canonical subgraph  $G'$  which is a copy of  $K_{q,2l}$  and let  $\beta$  be the number of edge-colorings of  $G$  that contain a monochromatic subgraph  $G'$  which is a copy of  $K_{q,2l}$ . We will prove that both  $\alpha$  and  $\beta$  are at most  $(1/2)D^{(s+t)^2}$  and conclude by Claim 19 that the total number of colorings is at most  $\alpha + \beta \leq D^{(s+t)^2}$  as desired.

Let the vertex set of  $G' = K_{q,2l}$  be  $Q \sqcup R$ , where  $Q \in \binom{X}{q}$ ,  $R \in \binom{Y}{2l}$  and  $\{X, Y\} = \{S, T\}$ . Define  $a = |X|$  and  $b = |Y|$  so  $\{a, b\} = \{s, t\}$ .

### 6.2.1 The canonical case

Our goal is to show that  $\alpha \leq (1/2)D^{(s+t)^2}$ . Recall that for each  $x \in Q$ , the edges in  $E(x, R)$  all have the same color  $z_x$  which is called a *canonical color*. Let  $Z_c = \{z_x : x \in Q\}$  be the set of all canonical colors. For each edge  $xy$  with  $x \in Q, y \in Y \setminus (R \cup Z_c)$ , a color  $z_{xy} \neq z_x$  is called a *free color*. We will count the number of colorings of  $E(Q, Y)$ , and then remove  $Q$  to apply the induction hypothesis. For each coloring  $N_G$ , consider the following partition of  $Y \setminus (R \cup Z_c)$  into two parts:

$$\begin{aligned} Y_0 &= \{y \in Y \setminus (R \cup Z_c) : E(y, Q) \text{ sees at most } 11l - 1 \text{ free colors}\}, \\ Y_1 &= \{y \in Y \setminus (R \cup Z_c) : E(y, Q) \text{ sees at least } 11l \text{ free colors}\}. \end{aligned}$$

We claim that the length of strongly rainbow colored paths that lie between  $Q$  and  $Y_1$  is bounded.

**Claim 20** *If there exists a strongly rainbow colored path  $P = P_{2l-2} \subset E(Q, Y_1)$  with both end-vertices  $u, v \in Q$ , then there exists a  $C_{2l}$  in  $G^*$ .*

*Proof of Claim 20.* Clearly,  $P$  extends to a linear  $P_{2l-2}$  in  $G^*$ . We may assume both  $z_u, z_v \notin V(P^*)$ , where  $P^* = \{e \cup \{z_e\} : e \in P\}$  is the extension of  $P$ . Otherwise, suppose w.l.o.g.  $z_u \in V(P^*)$ , let  $y$  be the vertex next to  $u$  in  $P$ , let  $S_y$  be of maximum size among sets

$$\{x \in Q : xy \text{ all colored by distinct free colors}\}.$$

Since  $y \in Y_1$ ,  $|S_y| \geq 11l$ . Note that  $|V(P^*)| = 4l - 3$  and  $|V(P^*) \cap Y_1| \geq l - 1$ , we have  $|S_y \setminus V(P^*)| \geq 11l - (4l - 3 - (l - 1)) \geq 8l$ . Since  $|V(P^*)| < 4l$ ,  $E(y, S_y)$  is rainbow, and  $G'$  is  $Q$ -canonical, there must be at least  $4l$  vertices in  $S_y \setminus V(P^*)$  whose canonical color is not in  $V(P^*)$ . Among these  $4l$  vertices there is at least one  $u'$  with  $z_{u'y} \notin V(P^*)$ . Replacing  $y$  by  $u'$ , we get a strongly rainbow colored path of length  $2l - 2$  with  $z_u \notin V(P^*)$ .

Now, Since  $|R| = 2l$ , we can find a vertex  $y \in R$  such that  $y \notin \{z_e : e \in P\}$ . Further, since both  $z_u, z_v \notin V(P^*)$  and  $z_u \neq z_v$ , the set of edges

$$P^* \cup \{uyz_u, vyz_v\}$$

forms a copy of  $C_{2l}$  in  $G^*$ . □

Thanks to this observation about strongly rainbow paths, we can bound the number of colorings on  $E(Q, Y_1)$  as follows. It is convenient to use the following notation.

**Definition 21** Given  $X' \subset X$  and  $Y' \subset Y$ , let  $\#E(X', Y')$  be the number of ways to color the edges in  $E(X', Y')$ .

**Claim 22**  $\#E(Q, Y_1) \leq (2l)^q \cdot (32l^2)^{bq} \cdot (qb)^{2lq} \cdot \sigma^{6lq+8l^2b}$ .

*Proof of Claim 22.* By Claim 20, according to the length of the longest strongly rainbow colored path starting at a vertex,  $Q$  can be partitioned into  $2l - 3$  parts  $\bigsqcup_{i=1}^{2l-3} Q_i$ , where

$$Q_i = \{x \in Q : \text{the longest strongly rainbow colored path starting at } x \text{ and contained in } E(Q, Y_1) \text{ has length } i\}.$$

For each  $i$ , let  $q_i = |Q_i|$ . We now bound the number of colorings of the edges in  $E(Q_i, Y_1)$ .

Firstly, for each  $x \in Q_i$ , choose an  $i$ -path  $P_x \subset E(Q, Y_1)$  starting at  $x$  and color it strongly rainbow. The number of ways to choose and color these paths for all the vertices  $x \in Q_i$  is at most

$$((qb)^{\lceil (i+1)/2 \rceil} \sigma^i)^{q_i} \leq (qb\sigma)^{iq_i}.$$

Fix an  $x \in Q_i$ . Partition  $Y_1$  into 3 parts depending on whether  $y$  is on the extension  $P_x^*$  of the path starting at  $x$ , or the color of  $xy$  is on  $P_x^*$  or else, i.e.  $Y_1 = \bigsqcup_{j=1}^3 Y_{i,x}^{(j)}$ , where

$$\begin{aligned} Y_{i,x}^{(1)} &= Y_1 \cap V(P_x^*), \\ Y_{i,x}^{(2)} &= \{y \in Y_1 \setminus Y_{i,x}^{(1)} : z_{xy} \in V(P_x^*)\}, \\ Y_{i,x}^{(3)} &= Y_1 \setminus (Y_{i,x}^{(1)} \cup Y_{i,x}^{(2)}). \end{aligned}$$

Depending on the part of  $Y_1$  that a vertex  $y$  lies in, we can get different restrictions on the coloring of the edges in  $E(y, Q_i)$ .

- If  $y \in Y_{i,x}^{(1)}$ , then  $z_{xy}$  has as many as  $\sigma$  choices. Note that  $|P_x^*| = 2i + 1$ , and  $|Y_{i,x}^{(1)}| \leq i + \lceil i/2 \rceil \leq 2i$ . This gives  $\#E(x, Y_{i,x}^{(1)}) \leq \sigma^{2i}$ .
- If  $y \in Y_{i,x}^{(2)}$ , then  $z_{xy} \in V(P_x^*)$ , so there are at most  $2i + 1$  choices for this color and  $\#E(x, Y_{i,x}^{(2)}) \leq (2i + 1)^b$ .
- Lastly, let  $|Y_{i,x}^{(3)}| = b_{i,x}$ . If  $y \in Y_{i,x}^{(3)}$ , then  $xy$  extends  $P_x$  into a strongly rainbow colored path  $P'_x = P_x \cup \{xy\}$  of length  $i + 1$ , which forces the edges  $x'y$  to be colored by  $V(P'_x)$  for each  $x' \in Q_i \setminus V(P'_x)$ . Otherwise, the path  $P'_x \cup \{x'y\}$  is a strongly rainbow colored path of length  $i + 2$  starting at a vertex  $x' \in Q_i$ , contradicting the definition of  $Q_i$ . Therefore,  $z_{x'y}$  has at most  $2i + 3$  choices if  $x' \in Q_i \setminus V(P'_x)$ . Putting this together, for each  $y \in Y_{i,x}^{(3)}$ , we have

$$\#E(Q_i \setminus V(P'_x), y) \leq (2i + 3)^{q_i}.$$

Noticing that  $|Q_i \cap V(P'_x)| \leq i + 1 + \lceil (i + 1)/2 \rceil \leq 2i + 1$ , we have

$$\#E(Q_i, y) \leq \#E(Q_i \cap V(P'_x), y) \cdot \#E(Q_i \setminus V(P'_x), y) \leq \sigma^{2i+1} (2i + 3)^{q_i}.$$

Hence the number of ways to color  $E(x, Y_1) \cup E(Q_i, Y_{i,x}^{(3)})$  is at most

$$2^b \cdot \sigma^{2i} \cdot (2i + 1)^b \cdot \sigma^{(2i+1)b_{i,x}} (2i + 3)^{q_i b_{i,x}}.$$

The term  $2^b$  arises above since  $Y_{i,x}^{(1)}$  has already been fixed before this step, so we just need to partition  $Y_1 \setminus Y_{i,x}^{(1)}$  to get  $Y_{i,x}^{(2)}$  and  $Y_{i,x}^{(3)}$ .

Now we remove  $x$  from  $Q_i$ ,  $Y_{i,x}^{(3)}$  from  $Y_1$  and repeat the above steps until we have the entire  $E(Q_i, Y_1)$  colored. Note that  $\sum_{x \in Q_i} b_{i,x} \leq b$ , and that  $i \leq 2l - 3$  which implies  $2i + 3 < 4l$ . We obtain

$$\begin{aligned} \#E(Q_i, Y_1) &\leq (qb\sigma)^{iq_i} \prod_{x \in Q_i} 2^b \cdot \sigma^{2i} \cdot (2i + 1)^b \cdot \sigma^{(2i+1)b_{i,x}} (2i + 3)^{q_i b_{i,x}} \\ &\leq (qb\sigma)^{2lq_i} \prod_{x \in Q_i} 2^b \cdot \sigma^{4l+4lb_{i,x}} \cdot (4l)^{b+q_i b_{i,x}} \\ &\leq (qb\sigma)^{2lq_i} \cdot 2^{bq_i} \cdot \sigma^{4lq_i+4lb} \cdot (4l)^{bq_i+bq_i} \\ &= (32l^2)^{bq_i} \cdot (qb)^{2lq_i} \cdot \sigma^{6lq_i+4lb}. \end{aligned}$$

Because  $\sum_{i=1}^{2l-3} q_i = q$ , taking the product over  $i \in [2l - 3]$ , we obtain

$$\begin{aligned} \#E(Q, Y_1) &\leq (2l - 3)^q \prod_{i=1}^{2l-3} \#E(Q_i, Y_1) \leq (2l - 3)^q \prod_{i=1}^{2l-3} (32l^2)^{bq_i} \cdot (qb)^{2lq_i} \cdot \sigma^{6lq_i+4lb} \\ &\leq (2l)^q \cdot (32l^2)^{bq} \cdot (qb)^{2lq} \cdot \sigma^{6lq+8l^2b}, \end{aligned}$$

where  $(2l - 3)^q$  counts the number of partitions of  $Q$  into the  $Q_i$ . □

Since  $G' = E(Q, R)$  is  $Q$ -canonical,

$$\#E(Q, R) \leq \sigma^q.$$

As  $|Z_c| \leq q$ ,

$$\#E(Q, Y \cap Z_c) \leq \sigma^{q^2}.$$

By definition of  $Y_0$ ,

$$\#E(Q, Y_0) \leq (\sigma^{11l}(11l+1)^q)^b \leq (\sigma^{11l}(12l)^q)^b.$$

Therefore to color  $E(Q, Y)$ , we need to first choose the subsets  $R$  and  $Z_c \cap Y$  of  $Y$  and then take a partition to get  $Y_0$  and  $Y_1$ . We color each of  $E(Q, R)$ ,  $E(Q, Y \cap Z_c)$ ,  $E(Q, Y_0)$  and  $E(Q, Y_1)$ . This gives

$$\begin{aligned} \#E(Q, Y) &\leq b^{2l} b^q 2^b \cdot \#E(Q, R) \cdot \#E(Q, Y \cap Z_c) \cdot \#E(Q, Y_0) \cdot \#E(Q, Y_1) \\ &\leq b^{2l} b^q 2^b \cdot \sigma^q \cdot \sigma^{q^2} \cdot (\sigma^{11l}(12l)^q)^b \cdot [(2l)^q \cdot (32l^2)^{bq} \cdot (qb)^{2lq} \cdot \sigma^{6lq+8l^2b}] \\ &= b^{2l} 2^b \cdot (2lb)^q \cdot (384l^3)^{qb} \cdot (qb)^{2lq} \cdot \sigma^{q+q^2+11lb+6lq+8l^2b}. \end{aligned}$$

Finally, we apply the induction hypothesis to count the number ways to color  $G = K_{s,t}$ . Recall that  $q = s^\epsilon < s/\log s < t \leq s$ ,  $\sigma \leq 2l(s+t) \leq 4ls$ . There are two cases.

- $(X, Y) = (S, T)$  **and**  $(a, b) = (s, t)$

Recall that we must first choose  $Q \subset X$ .

$$\begin{aligned} \#E(X, Y) &\leq s^q \cdot \#E(Q, Y) \cdot \#E(X \setminus Q, Y) \\ &\leq s^q \cdot t^{2l} 2^t \cdot (2lt)^q \cdot (384l^3)^{qt} \cdot (qt)^{2lq} \cdot \sigma^{q+q^2+11lt+6lq+8l^2t} \cdot D^{(s+t-q)^2} \\ &\leq s^q t^{2l} \cdot 2^t \cdot (2lt)^q \cdot (384l^3)^{qt} \cdot (qt)^{2lq} \cdot (4ls)^{q+q^2+11lt+6lq+8l^2t} \cdot D^{(s+t-q)^2} \\ &\leq s^q t^{2l} \cdot t^q \cdot (2^{1/q}(2l)^{1/t} 384l^3)^{qt} \cdot (qt)^{2lq} \cdot (4ls)^{9l^2t} \cdot D^{-2qt} \cdot D^{-2qs+q^2} \cdot D^{(s+t)^2} \\ &\leq \left( \frac{2^{1/q}(2l)^{1/t} 384l^3 (4l)^{9l^2/q}}{D^2} \right)^{qt} \cdot t^{2l} (st)^q (qt)^{2lq} s^{9l^2t} \cdot D^{-qs} \cdot D^{(s+t)^2} \\ &\leq \left( \frac{2^{1/q}(2l)^{1/t} 384l^3 (4l)^{9l^2/q}}{D^2} \right)^{qt} \cdot \frac{q^{2lq} t^{2l+q+2lq} s^{q+9l^2t}}{D^{qs}} \cdot D^{(s+t)^2} \\ &\leq \frac{1}{4} D^{(s+t)^2}. \end{aligned}$$

To show the last inequality above, it is obvious that  $2^{1/q}(2l)^{1/t} 384l^3 (4l)^{9l^2/q} < D^2$  for large enough  $D$ , so we are left to show that  $4q^{2lq} t^{2l+q+2lq} s^{q+9l^2t} < D^{qs}$  for large  $D$ . Taking logarithms, we have

$$\begin{aligned} \log \left( 4q^{2lq} t^{2l+q+2lq} s^{q+9l^2t} \right) &= 2 + 2lq \log q + (2l+q+2lq) \log t + (q+9l^2t) \log s \\ &\leq qs \log D. \end{aligned}$$

This holds for large  $D$  because  $qs$  has the dominating growth rate among all the above terms as  $q = s^\epsilon$  and  $s$  is large.

- $(X, Y) = (T, S)$  and  $(a, b) = (t, s)$

$$\begin{aligned}
\#E(X, Y) &\leq t^q \cdot \#E(Q, X) \cdot \#E(Y \setminus Q, X) \\
&\leq t^q \cdot s^{2l} 2^s \cdot (2ls)^q \cdot (384l^3)^{qs} \cdot (qs)^{2lq} \cdot \sigma^{q+q^2+11ls+6lq+8l^2s} \cdot D^{(s+t-q)^2} \\
&\leq t^q s^{2l} \cdot 2^s \cdot (2ls)^q \cdot (384l^3)^{qs} \cdot (qs)^{2lq} \cdot (4ls)^{q+q^2+11ls+6lq+8l^2s} \cdot D^{(s+t-q)^2} \\
&\leq t^q s^{2l} \cdot s^q \cdot (2^{1/q} (2l)^{1/s} 384l^3)^{qs} \cdot (qs)^{2lq} \cdot (4ls)^{9l^2s} \cdot D^{-2qs} \cdot D^{-2qt+q^2} \cdot D^{(s+t)^2} \\
&\leq \left( \frac{2^{1/q} (2l)^{1/s} 384l^3 (4l)^{9l^2/q}}{D^2} \right)^{qs} \cdot s^{2l} (st)^q (qs)^{2lq} s^{9l^2s} \cdot D^{-qt} \cdot D^{(s+t)^2} \\
&\leq \left( \frac{2^{1/q} (2l)^{1/s} 384l^3 (4l)^{9l^2/q}}{D^2} \right)^{qs} \cdot \frac{q^{2lq} t^q s^{2l+q+2lq+9l^2t}}{D^{qt}} \cdot D^{(s+t)^2} \\
&\leq \frac{1}{4} D^{(s+t)^2}.
\end{aligned}$$

Again, to show the last inequality above, it is clear that  $2^{1/q} (2l)^{1/s} 384l^3 (4l)^{9l^2/q} < D^2$  for large  $D$ , so we are left to show that  $4q^{2lq} t^q s^{2l+q+2lq+9l^2s} < D^{qs}$  for large  $D$ . Taking logarithms, we have

$$\begin{aligned}
\log \left( 4q^{2lq} t^q s^{2l+q+2lq+9l^2s} \right) &= 2 + 2lq \log q + q \log t + (2l + q + 2lq + 9l^2s) \log s \\
&\leq qt \log D.
\end{aligned}$$

This holds for large  $D$  because  $qt$  has the dominating growth rate among all the above terms.

In summary, the number of colorings of  $G$  such that there exists a  $G' \subset G$  that is a  $Q$ -canonical  $K_{q,2l}$  is

$$\alpha \leq \frac{1}{4} D^{(s+t)^2} + \frac{1}{4} D^{(s+t)^2} = \frac{1}{2} D^{(s+t)^2}.$$

### 6.2.2 The monochromatic case

Our goal is to show that  $\beta \leq (1/2) D^{(s+t)^2}$ . Recall that the vertex set of  $G' = K_{q,2l}$  is  $Q \sqcup R$ , where  $Q \in \binom{X}{q}$  and  $R \in \binom{Y}{2l}$ . The term *canonical color* now refers to the only color  $z_c$  that is used to color all edges of  $G'$ , and  $Z_c = \{z_c\}$  still means the set of canonical colors. A *free color* is a color that is not  $z_c$ . As before we will count the number of colorings of  $E(Q, Y)$ , and then remove  $Q$  to apply the induction hypothesis.

Let  $Y_1 = Y \setminus (R \cup Z_c)$ . Similar to Claim 20, we claim that the length of a strongly rainbow colored path between  $Q$  and  $Y_1$  is bounded.

**Claim 23** *If there exists a strongly rainbow colored path  $P = P_{4l-2} \subset E(Q, Y_1)$  with both end-vertices  $u, v \in Q$ , then there exists a  $C_{2l}$  in  $G^*$ .*

*Proof of Claim 23.* We observe that  $z_c$  appears in the path or the color of the path at most once, as  $P$  is strongly rainbow. Hence, by the pigeonhole principle, there exists a sub-path  $P'$  of length  $2l - 2$  such that  $z_c \notin V(P'^*)$  and both end-vertices  $u, v$  of  $P'$  are in  $Q$ .

Now, Since  $|R| = 2l$ , we can find two vertices  $y, y' \in R$  such that  $y, y' \notin \{z_e : e \in P'\}$ . Thus, the edges

$$P'^* \cup \{uyz_c, vy'z_c\}$$

yield a copy of  $C_{2l}$  in  $G^*$ . □

Again, we first use this claim to color  $E(Q, Y_1)$ .

**Claim 24**  $\#E(Q, Y_1) \leq (4l)^q \cdot (128l^2)^{qb} \cdot (qb)^{4lq} \cdot \sigma^{12lq+32l^2b}$ .

*Proof of Claim 24.* The proof proceeds exactly the same as that of Claim 22, except that  $Q$  is partitioned into  $4l - 3$  parts  $\bigsqcup_{i=1}^{4l-3} Q_i$ . So in the calculation at the end, we have  $i \leq 4l - 3$  which gives  $2i + 3 < 8l$  and

$$\begin{aligned} \#E(Q_i, Y_1) &\leq (qb\sigma)^{iq_i} \prod_{x \in Q_i} 2^b \cdot \sigma^{2i} \cdot (2i + 1)^b \cdot \sigma^{(2i+1)b_{i,x}} (2i + 3)^{q_i b_{i,x}} \\ &\leq (qb\sigma)^{4lq_i} \prod_{x \in Q_i} 2^b \cdot \sigma^{8l+8lb_{i,x}} \cdot (8l)^{b+q_i b_{i,x}} \\ &\leq (qb\sigma)^{4lq_i} \cdot 2^{bq_i} \cdot \sigma^{8lq_i+8lb} \cdot (8l)^{bq_i+bq_i} \\ &\leq (128l^2)^{bq_i} \cdot (qb)^{4lq_i} \cdot \sigma^{12lq_i+8lb}. \end{aligned}$$

Again, note that  $\sum_{i=1}^{4l-3} q_i = q$ . Taking the product over  $i \in [4l - 3]$ , we obtain

$$\begin{aligned} \#E(Q, Y_1) &\leq (4l - 3)^q \prod_{i=1}^{4l-3} \#E(Q_i, Y_1) \leq (4l - 3)^q \prod_{i=1}^{4l-3} (128l^2)^{bq_i} \cdot (qb)^{4lq_i} \cdot \sigma^{12lq_i+8lb} \\ &\leq (4l)^q \cdot (128l^2)^{qb} \cdot (qb)^{4lq} \cdot \sigma^{12lq+32l^2b}, \end{aligned}$$

where  $(4l - 3)^q$  counts the number of partitions of  $Q$  into the  $Q_i$ . □

Similarly, to color  $E(Q, Y)$ , we need to choose the subsets  $R$  and  $Y \cap Z_c$ , and what remains is  $Y_1$ . Consequently,

$$\begin{aligned} \#E(Q, Y) &\leq b^{2l}b \cdot \#E(Q, R) \cdot \#E(Q, Y \cap Z_c) \cdot \#E(Q, Y_1) \\ &\leq b^{2l}b \cdot \sigma \cdot \sigma^q \cdot [(4l)^q \cdot (128l^2)^{qb} \cdot (qb)^{4lq} \cdot \sigma^{12lq+32l^2b}] \\ &= b^{2l+1}(4l)^q \cdot (128l^2)^{qb} \cdot (qb)^{4lq} \cdot \sigma^{1+q+12lq+32l^2b}. \end{aligned}$$

We apply the induction hypothesis to count the number ways to color  $G = K_{s,t}$ . Recall that  $q = s^\epsilon < s/\log s < t \leq s$ ,  $\sigma \leq 2l(s + t) \leq 4ls$ . We split the calculation into two cases.

- $(X, Y) = (S, T)$  and  $(a, b) = (s, t)$

Recall that we must first choose  $Q \subset X$ .

$$\begin{aligned}
\#E(X, Y) &\leq s^q \cdot \#E(Q, Y) \cdot \#E(X \setminus Q, Y) \\
&\leq s^q \cdot t^{2l+1} (4l)^q \cdot (128l^2)^{qt} \cdot (qt)^{4lq} \cdot \sigma^{1+q+12lq+32l^2t} \cdot D^{(s+t-q)^2} \\
&\leq s^q t^{2l+1} \cdot (4l)^q \cdot (128l^2)^{qt} \cdot (qt)^{4lq} \cdot (4ls)^{1+q+12lq+32l^2t} \cdot D^{(s+t-q)^2} \\
&\leq s^q t^{2l+1} \cdot ((4l)^{1/t} 128l^2)^{qt} \cdot (qt)^{4lq} \cdot (4ls)^{33l^2t} \cdot D^{-2qt} \cdot D^{-2qs+q^2} \cdot D^{(s+t)^2} \\
&\leq \left( \frac{(4l)^{1/t} 128l^2 (4l)^{33l^2/q}}{D^2} \right)^{qt} \cdot t^{2l+1} s^q (qt)^{4lq} s^{33l^2t} \cdot D^{-qs} \cdot D^{(s+t)^2} \\
&\leq \left( \frac{(4l)^{1/t} 128l^2 (4l)^{33l^2/q}}{D^2} \right)^{qt} \cdot \frac{q^{4lq} t^{2l+1+4lq} s^{q+33l^2t}}{D^{qs}} \cdot D^{(s+t)^2} \\
&\leq \frac{1}{4} D^{(s+t)^2}.
\end{aligned}$$

To show the last inequality above, it is obvious to see that  $(4l)^{1/t} 128l^2 (4l)^{33l^2/q} < D^2$  for large  $D$ , so we are left to show that  $4q^{4lq} t^{2l+1+4lq} s^{q+33l^2t} < D^{qs}$  for large  $D$ . Taking logarithms, we have

$$\begin{aligned}
\log \left( 4q^{4lq} t^{2l+1+4lq} s^{q+33l^2t} \right) &= 2 + 4lq \log q + (2l + 1 + 4lq) \log t + (q + 33l^2t) \log s \\
&\leq qs \log D.
\end{aligned}$$

This holds for large  $D$  because  $qs$  has dominating growth rate among all the terms above.

- $(X, Y) = (T, S)$  and  $(a, b) = (t, s)$

$$\begin{aligned}
\#E(X, Y) &\leq t^q \cdot \#E(Q, Y) \cdot \#E(X \setminus Q, Y) \\
&\leq t^q \cdot s^{2l+1} (4l)^q \cdot (128l^2)^{qs} \cdot (qs)^{4lq} \cdot \sigma^{1+q+12lq+32l^2s} \cdot D^{(s+t-q)^2} \\
&\leq t^q s^{2l+1} \cdot (4l)^q \cdot (128l^2)^{qs} \cdot (qs)^{4lq} \cdot (4ls)^{1+q+12lq+32l^2s} \cdot D^{(s+t-q)^2} \\
&\leq t^q s^{2l+1} \cdot ((4l)^{1/s} 128l^2)^{qs} \cdot (qs)^{4lq} \cdot (4ls)^{33l^2s} \cdot D^{-2qs} \cdot D^{-2qt+q^2} \cdot D^{(s+t)^2} \\
&\leq \left( \frac{(4l)^{1/s} 128l^2 (4l)^{33l^2/q}}{D^2} \right)^{qs} \cdot s^{2l+1} t^q (qs)^{4lq} s^{33l^2s} \cdot D^{-qt} \cdot D^{(s+t)^2} \\
&\leq \left( \frac{(4l)^{1/s} 128l^2 (4l)^{33l^2/q}}{D^2} \right)^{qs} \cdot \frac{q^{4lq} t^q s^{2l+1+4lq+33l^2t}}{D^{qt}} \cdot D^{(s+t)^2} \\
&\leq \frac{1}{4} D^{(s+t)^2}.
\end{aligned}$$

To show the last inequality above, it is obvious to see that  $(4l)^{1/s} 128l^2 (4l)^{33l^2/q} < D^2$  for large  $D$ , so we are left to show that  $4q^{4lq} t^q s^{2l+1+4lq+33l^2t} < D^{qt}$  for large  $D$ . Taking logarithms, we have

$$\begin{aligned}
\log \left( 4q^{4lq} t^q s^{2l+1+4lq+33l^2t} \right) &= 2 + 4lq \log q + q \log t + (2l + 1 + 4lq + 33l^2t) \log s \\
&\leq qt \log D.
\end{aligned}$$

This holds for large  $D$  because  $qt$  has the dominating growth rate among all the terms above.

Finally, we have

$$\beta \leq \frac{1}{4}D^{(s+t)^2} + \frac{1}{4}D^{(s+t)^2} = \frac{1}{2}D^{(s+t)^2},$$

and the proof is complete.  $\square$

## 7 Concluding remarks

- A straightforward corollary of Theorem 2 is the very same result for hypergraph paths  $P_k$ . Indeed, for the upper bound on  $\text{Forb}_r(n, P_k)$  one has to just observe that  $P_k \subset C_{2\lceil(k+1)/2\rceil}$ , while the lower bound is trivial.
- Although we were unable to use hypergraph containers to prove Theorem 2 it would be very interesting to give a new proof using containers. In particular, this would entail proving some supersaturation type results for this problem which may be of independent interest. It would also likely yield some further results in the random setting which we have not addressed.
- The main open problem raised by our work is to solve the analogous question for larger  $r$  and for odd cycles (Conjecture 3).

For  $r = 3$ , our method will not work for odd cycles as it relies on finding a bipartite structure from which it is difficult to extract odd 3-uniform cycles (although this technical hurdle could be overcome to solve the corresponding extremal problem in [37]).

For larger  $r$ , our method does not work because the cost of decomposing a complete  $r$ -graph into complete  $r$ -partite subgraphs is too large to remain an error term. More precisely, for  $r = 3$ , we implicitly applied Lemma 12 (in the proof of Lemma 8) to reduce the number of ways to color a graph to at most  $2^{O(n^2)}$  instead of the trivial  $2^{O(n^2 \log \log n)}$ . But for  $r > 3$  the main term in the calculation turns out to be  $2^{O(n^{r-1}(\log n)^{(r-3)/(r-2)})}$  which comes from choosing the colors for the copies of  $K_{s_i:r-1}$  (see Section 3). This cannot be improved due to Theorem 6 and Lemma 10 each of which gives a bound that is sharp in order of magnitude. Consequently, even if we proved a version of Lemma 12 for  $r > 3$  (and the tools we have developed should suffice to provide such a proof) this would not improve Theorem 2 for  $r > 3$ .

- Another way to generalize the result of Morris-Saxton to hypergraphs is to consider similar enumeration questions when the underlying  $r$ -graph is linear, meaning that every two edges share at most one vertex. Here the extremal results have recently been proved in [20] and the formulas are similar to the case of graphs. The special case of this question for linear triple systems without a  $C_3$  is related to the Ruzsa-Szemerédi  $(6, 3)$  theorem and sets without 3-term arithmetic progressions.

**Acknowledgment.** We are very grateful to Rob Morris for clarifying some technical parts of the proof in [42] at the early stages of this project. After those discussions, we realized



that the method of hypergraph containers would not apply easily to prove Theorem 2 and we therefore developed new ideas. We are also grateful to Jozsef Balogh for providing us with some pertinent references, and to Jie Han for pointing out that our proof of Theorem 2 applies for  $r > 3$  and odd  $k$ .

## References

- [1] N. Alon, J. Balogh, B. Bollobás and R. Morris, The structure of almost all graphs in a hereditary property, *J. Combin. Theory, Ser. B* 101 (2011), 85–110.
- [2] B. Bollobás, Cycles modulo  $k$ , *Bull. London Math. Soc.* 9 (1977), 97–98.
- [3] S. Bublitz, Decomposition of graphs and monotone formula size of homogeneous functions, *Acta Informatica* 23 (1986), 689–696.
- [4] J. Balogh, T. Bohman and D. Mubayi, Erdős-Ko-Rado in random hypergraphs, *Combin. Probab. Comput.* 18 (2009), 629–646.
- [5] B. Bollobás, B. Narayanan and A. Raigorodskii, On the stability of the Erdős-Ko-Rado theorem, arXiv:1408.1288.
- [6] J. Balogh, B. Bollobás and M. Simonovits, On the number of graphs without forbidden subgraph, *J. Combin. Theory, Ser. B* 91 (2004), 1–24.
- [7] ———, The typical structure of graphs without given excluded subgraphs, *Random Structures Algorithms* 34 (2009), no. 3, 305–318.
- [8] ———, The fine structure of octahedron-free graphs, *J. Combin. Theory Ser. B* 101 (2011), no. 2, 67–84.
- [9] J. Balogh, S. Das, M. Delcourt, H. Liu, M. Sharifzadeh, Intersecting families of discrete structures are typically trivial, *J. Combin. Theory, Ser. A* 132 (2015), 224–245.
- [10] J. Balogh and D. Mubayi, Almost all triple systems with independent neighborhoods are semi-bipartite, *J. Combin. Theory Ser. A* 118 (2011), no. 4, 1494–1518.
- [11] ———, Almost all triangle-free triple systems are tripartite, *Combinatorica* 32 (2012), no. 2, 143–169.
- [12] J. Balogh and A.Z. Wagner, On the number of union-free families, accepted, *Israel J. Math.*
- [13] J. Balogh, R. Morris and W. Samotij, Independent sets in hypergraphs, *J. Amer. Math. Soc.* 28 (2015), 669–709.
- [14] J. Balogh and W. Samotij, The number of  $K_{m,m}$ -free graphs, *Combinatorica* 31 (2011), 131–150.
- [15] ———, The number of  $K_{s,t}$ -free graphs, *J. London Math. Soc.* 83 (2011), 368–388.

- [16] J. A. Bondy and M. Simonovits, Cycles of even length in graphs, *J. Combin. Theory Ser. B* 16 (1974), 97–105.
- [17] G. Brightwell and S. Goodall, The number of partial orders of fixed width, *Order* 13 (1996), no. 4, 315–337.
- [18] G. Brightwell, D.A. Grable and H. J. Prömel, Forbidden induced partial orders, *Discrete Math.* 201 (1999), no. 1-3, 53–80.
- [19] F.R.K. Chung, P. Erdős and J. Spencer, On the decomposition of graphs into complete bipartite subgraphs *Studies in Pure Math.*(1983), pp. 95–101 Mem. of P. Turán
- [20] C. Collier-Cartaino, N. Graber and T. Jiang, Linear Turán numbers of  $r$ -uniform linear cycles and related Ramsey numbers, arXiv:1404.5015v2, preprint, 2014.
- [21] D. Conlon and W.T. Gowers, Combinatorial theorems in sparse random sets, *Ann. of Math.* 184 (2016), 367–454.
- [22] S. Das, T. Tran, Removal and Stability for Erdős-Ko-Rado *SIAM J. Discrete Math.* 30 (2), 1102–1114.
- [23] P. Devlin, J. Kahn, On "stability" in the Erdős-Ko-Rado theorem, arXiv:1502.05692
- [24] A. Diwan, Cycles of even lengths modulo  $k$ , *J. Graph Theory* 65 no. 3 (2010), 246–252.
- [25] P. Erdős, D.J. Kleitman and B.L. Rothschild, Asymptotic enumeration of  $K_n$ -free graphs, in *Colloquio Internazionale sulle Teorie Combinatorie* (Rome, 1973), Vol. II, pp. 19–27. *Atti dei Convegni Lincei* **17**, Accad. Naz. Lincei, Rome, 1976.
- [26] P. Erdős, P. Frankl and V. Rödl, The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent, *Graphs and Combinatorics* **2** (1986), 113–121.
- [27] P. Erdős, On extremal problems of graphs and generalized graphs, *Israel J. Math.* 2 (1964), 183–190.
- [28] P. Erdős and A.H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.* 52 (12) (1946), 1087–1091.
- [29] Z. Füredi, T. Jiang, Hypergraph Turán numbers of linear cycles, *J. Combin. Theory Ser. A* 123 (1) (2014) 252–270.
- [30] Z. Füredi, T. Jiang and R. Seiver, Exact solution of the hypergraph Turán problem for  $k$ -uniform linear paths, *Combinatorica* 34 (2014), 299–322.
- [31] C. Hundack, H.J. Prömel and Angelika Steger, Extremal graph problems for graphs with a color-critical vertex, *Combin. Probab. Comput.* 2 (1993), no. 4, 465–477.
- [32] D. Kleitman and K. Winston, On the number of graphs without 4-cycles, *Discrete Math.* **41** (1982), 167–172.

- [33] D. Kleitman and D. Wilson, On the number of graphs which lack small cycles, manuscript, 1996.
- [34] Ph.G. Kolaitis, H.J. Prömel and B.L. Rothschild,  $K_{l+1}$ -free graphs: asymptotic structure and a 0-1 law, *Transactions of the Amer. Math. Soc.* Vol. 303, No. 2 (Oct., 1987), pp. 637–671
- [35] ———, Asymptotic enumeration of partial orders on a finite set, *Trans. Amer. Math. Soc.* 205 (1975), 205–220.
- [36] T. Kővári, V. Sós and P. Turán, On a problem of K. Zarankiewicz, *Colloquium Math.* 3 (1954), 50–57.
- [37] A. Kostochka, D. Mubayi and J. Verstraëte, Turán problems and shadows I: paths and cycles, *J. Combin. Theory, series A* 129 (2015), 57–79.
- [38] ———, Turán problems and shadows II: trees, *J. Combin. Theory Ser. B* 122 (2017), 457–478.
- [39] ———, Turán problems and shadows III: expansions of graphs, *SIAM J. on Discrete Math.* Vol. 29, No. 2 (2015), pp. 868–876
- [40] E. Lamken and B.L. Rothschild, The numbers of odd-cycle-free graphs, Finite and infinite sets, Vol. I, II (Eger, 1981), *Colloq. Math. Soc. János Bolyai* vol. 37, North-Holland, Amsterdam, 1984, pp. 547–553.
- [41] C.-H. Liu, J. Ma, Cycle lengths and minimum degree of graphs, arXiv:1508.07912, preprint, 2015.
- [42] R. Morris and D. Saxton, The number of  $C_{2l}$ -free graphs, *Adv. Math.* 298 (2016), 534–580.
- [43] D. Mubayi and G. Turan, Finding bipartite subgraphs efficiently, *Information Processing Letters* Volume 110, Issue 5, 1 (2010), 174–177.
- [44] B. Nagle and V. Rödl, The asymptotic number of triple systems not containing a fixed one, *Discrete Math.* 235 (2001), no. 1-3, 271–290, *Combinatorics* (Prague, 1998).
- [45] B. Nagle, V. Rödl and M. Schacht, Extremal hypergraph problems and the regularity method, Topics in Discrete Mathematics, *Algorithms Combin.* **26** (2006), 247–278.
- [46] D. Osthus, D. Kühn, T. Townsend, and Y. Zhao, On the structure of oriented graphs and digraphs with forbidden tournaments or cycles, arXiv:1404.6178, preprint, 2014.
- [47] Y. Person and M. Schacht, Almost all hypergraphs without Fano planes are bipartite, *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, (2009), pp. 217–226.
- [48] H.J. Prömel and A. Steger, The asymptotic number of graphs not containing a fixed color-critical subgraph, *Combinatorica* 12 (1992), no. 4, 463–473.

- [49] R.W. Robinson, Counting labeled acyclic digraphs, *New directions in the theory of graphs* (Proc. Third Ann Arbor Conf., Univ. Michigan, Ann Arbor, Mich., 1971) (1973), pp. 239–273.
- [50] D. Saxton and A. Thomason, Hypergraph containers, *Inventiones Math.* **201** (2015), 1–68.
- [51] R.P. Stanley, Acyclic orientations of graphs, *Discrete Math.* 5 (1973), 171–178.
- [52] B. Sudakov and J. Verstraëte, The extremal function for cycles of length  $l \bmod k$ , *Elec. J. of Combin.* **22** (2015), #P00 .
- [53] C. Thomassen, Graph decomposition with applications to subdivisions and path systems modulo  $k$ , *J. Graph Theory* 7 (1983), 261–271.
- [54] P. Turán, On an extremal problem in graph theory, *Matematikai és Fizikai Lapok* (in Hungarian) 48 (1941), 436–452.
- [55] Z. Tuza, Covering of graphs by complete bipartite subgraphs; complexity of 0-1 matrices, *Combinatorica* 4 (1) (1984), 111–116.